9-2019

Digital Collection Assessment and Use

Tammy Troup
Bucknell University, tlt014@bucknell.edu

Recommended Citation
Troup, T. (2019). Digital Collection Assessment & Use. In #DLFTeach Toolkit: Lesson Plans for Digital Library Instruction (1st ed.). https://doi.org/10.21428/65a6243c.42628a94

Title

Assessing and Using Digital Collections

Session Specifics

Three hour workshop with two ten minute breaks

Instructional Partners

Discipline-specific teaching faculty

Audience

Mid-track undergraduates majoring in fields requiring a high-level of visual literacy about the cultural or physical world (anthropology, art, history, biology, geology, pre-medicine, etc.); graduate students; MLIS students learning about metadata development.

(Curricular) Context

This course was designed using a crosswalking method shared by Marcia Lei Zeng (Post-It note paper to tag, label, and simulate the crosswalking experience) and expanded upon in Metadata workshops held by the author.

The course supports the identification of metadata by individuals with local knowledge or subject-specific knowledge.

Learning Outcomes

The primary goal of this lesson is to familiarize students with the methods of metadata development and reuse, and to instill the confidence in their ability to contribute to curated knowledge. This is accomplished by following the following four steps:

1. Introduction to the effort required to create simple and robust metadata;
2. Use of the DPLA API to harvest metadata;
3. Practice of empathy in the assessment and use of digital collections, identify bias and how bias may be addressed, and identify gaps in access;
4. Enhancement of metadata for use by a specific user group, specifically through the identification of keywords, enhanced description, coverage, or additional fields.

Preparation

Audience --
Prior to workshop, students will need to identify a digital collection on the DPLA (https://dp.la/), which includes items affiliated with their area of special interest. Students should read "Queering the Catalog," [https://www.journals.uchicago.edu/doi/10.1086/669547#.W7JUM6XaoSQ.link] before workshop

Instructor --
Prepare printed images from the DPLA (https://dp.la/).

1. Images for students to discuss -- instructor can select any image shared with the DPLA, note that these image will be discussed in section 1.0 and cognitive biases will be assessed. Instructor should select images which could challenge student's identification of the content.
2. Maps to illustrate standards development (e.g.,
    a. hand drawn Florida map (few standards) (http://dc.lib.unc.edu/cdm/ref/collection/ncmaps/id/1097);
    b. engineer drawn (some standards) https://catalog.archives.gov/id/103396667;
    c. Shell Oil Company map (highly standardized) https://www.davidrumsey.com/luna/servlet/detail/RUMSEY~8~1~212056~5500200).

API practice **Note:** it is very important for the instructor to gain some experience working with the API prior to teaching the course, the DPLA documentation is excellent (https://pro.dp.la/developers/api-codex).
Sample datasets (backup).
Prepared bookmarks/open tabs/software Full List.

Materials

Computer with World Wide Web access, document software (Microsoft, Google, etc.), paper, pencils, Post-Its of various colors, DPLA API key (https://pro.dp.la/developers/policies#get-a-key), and supplement (https://drive.google.com/open?id=1FUTJWRNysj6hnzLe3W3kHccSfUmT8uqzbcwyMDOnr4Q)

Session Instructions/Steps

1.0 Developing Metadata

Introduction to the effort required to create simple and robust metadata

- Students --
    - **Organize** into small groups, each group **selects** a single image to discuss.
- Instructor --
    - **Introduce** the Panofsky-Shatford matrix (generics, specifics, abstracts)
    - **Discuss** cognitive biases and the effect on the transfer of information

- Students --
    - **Discuss** two or three cognitive biases and the potential effect of these biases on the process of creating metadata for the image in hand.
    - **Discuss** in small groups how the image could be “tagged” what “labels” could apply to the “tags,” what research may be required to identify more information, etc.
    - **Record** the labels and the tags on Post-Its
- Students --
    - Reconvene as a large group, and **crosswalk/map** all the groups labels/concepts
    - **Discuss** difficulties mapping labels
- Instructor --
    - **Introduce** the idea of standardized information (examples: hand drawn map, 19th century map, Rand-McNally map, Google map).
    - **Discuss** the ideas communicated through standardized information.
    - **Share and discuss** the standards funnel -- How is information created by humans? How is information discovered and used by humans? What information can be contributed? How is information crawled and used by machines?
- Students --
    - Use a **pro & con grid** (on paper or whiteboard) to list the benefits and drawbacks of standardized or centrally managed metadata.

2.0 Harvesting Metadata

Use DPLA API to harvest metadata

- Instructor --
    - Briefly **introduce** the idea of exposing metadata for use and reuse.
- Students --
    - **Discuss** possible ethical and legal issues to consider when using metadata created by others.
- Instructor --
    - **Demonstrate** harvest and creating calls
- Students --
    - **Construct** queries and **harvest** metadata DPLA API Instructions
    - **Collect** affiliated images from JSON results with an image scraper (e.g., Tab Save extension github.com/lmmx/tabsave)
- Students --
    - Use a **concept mapping** approach to describe the systems (machines, protocols, and people) used to harvest metadata

3.0 Assessing Metadata

Practice empathy in the assessment and use of digital collections, identify bias and how bias may be addressed, and identify gaps in access

- Students --
    - **Explore** the CMS Mukurtu

    - **Discuss** communities and attributes of communities (cultural, professional, knowledge domains, sub-Reddits, Twitter cultures, etc.)
- Instructor --
    - **Demonstrate** various tools and methods of metadata assessment **(word cloud guessing, sorting, visualization, etc.)**
- Students --
    - **Use** various tools and methods to investigate the metadata
- Instructor --
    - **Demonstrate** the assessment rubric
- Students --
    - **Investigate** a few original collections that metadata is drawn from, and use the rubric to **assess** the standards and policies related to metadata development.
- Students --
    - **Discuss and Identify** possible cultural bias, knowledge gaps, inclusive/exclusive approaches to the metadata
- Students --
    - **Identify** gaps in access -- Who may not discover this collection because some information is missing? Who is the primary contributor? Who is the primary audience of this collection? Who is excluded from sharing their knowledge? Are the labels useful for finding information?

4.0 Enhancing Metadata

Enhance metadata for use by a specific user group. Keywords, enhanced description, coverage, or additional fields

- Instructor --
    - **Discuss** sharing metadata with the collection creators and working collaboratively to enhance metadata.
- Students --
    - **Use** Criterion 3 of the assessment rubric to identify methods of making knowledge about the dataset more robust
- Instructor --
    - **Demonstrate** various tools and methods of refining and enhancing data (**Excel, OpenRefine,** or **Regex** depending on instructor's experience and comfort level).
- Students --
    - **Identify** methods of making access to the collection more robust, identify schemas, local elements, controlled vocabularies, local terms, which might be used.

Assessment

- Students identify issues which surround the curation of information by discussing biases and knowledge gaps.

- Students successfully use tools to collect and refine data.
- Students are empowered to research and to contribute to record-keeping.

Reflection

Metadata records are living records, which should be revisited and revised when new information is acquired or new perspectives require re-cataloging. As the myriad communities responsible for metadata creation, curation, and use continue to evolve, they need to consider best practices which allow for knowledge experts to collaborate and contribute. Perhaps we need interfaces which allow experts to fork and create enhanced collections. Whatever solutions are devised, subject experts should be involved as thinkers, contributors, and collaborators.